Trivial Similarity-Based Biases and Efforts to Avoid Bias in Courtroom Judgments

Research Thesis

Presented in partial fulfillment of the requirements for graduation *with research distinction* in Psychology in the undergraduate colleges of The Ohio State University

by

Sarah Thieken

The Ohio State University

April 2021

Project Advisor: Dr. Duane Wegener, Department of Psychology

Abstract

Previous research on jury trials has focused mainly on the effects of group similarities (the similarity-leniency hypothesis; Kerr et al., 1995) and ways to combat those biases to ensure fair deliberation. Previous research on the effects of trivial similarities shows increased liking and compliance towards individuals who share similarities, but the similarity can also lead to participants distancing themselves from the individual if they display negative characteristics (i.e., being rude). The current study investigated the effects of shared trivial similarities between juror and defendant and possible ways to reduce or possibly eliminate these effects using bias correction. The main hypothesis was that trivial similarities between juror and defendant influence the juror's ratings of the defendant, and when asked to correct, jurors will do so. For the design of the study, a single-session study was utilized where participants were randomly selected to either see a similar or non-similar defendant and provide ratings of guilt, fault, and responsibility. The bias correction instructions immediately followed the initial decisions in the same sitting and theories of bias were measured as a possible predictor of the shift in ratings from pre-to-post-correction instructions. Participants were 150 undergraduate psychology students at The Ohio State University who were participating for credit ranging from ages 18 to 37. The effects of similarity, correction instructions, and the interaction between the two all failed to reach significance. Theories of bias also did not significantly interact with condition to impact difference in perceptions but were trending in the direction I predicted. Specifically, post-correction guilt ratings were higher for participants who were in the similar condition which is trending in the expected direction. While there were limitations to this study including the possibility of the similarities being too subtle, if significant results were found, there would be implications for jury instructions in the courtroom and jury selection.

Introduction

In the United States, the criminal justice system is meant to be fair and just when punishing individuals. One of the main ways this is accomplished is by allowing a jury of peers to judge crimes and determine punishment. This right to a jury trial is guaranteed to each citizen in the 6th and 7th Amendments of the Bill of Rights (Cong. Rsch. Serv., n.d.). Although it has been prominent throughout our country's history, jury trials and the jury itself has gone through many changes to ensure that verdicts are fair. This has included expansion over the years in who can serve on a jury, including women, African Americans, and other minorities (each of whom had, at certain points in our country's history had been restricted in their jury service). When selecting jurors for a trial, every potential juror must go through jury selection where it is determined whether they are going to be fair during the trial. If there is an identified potential for bias, jurors can be removed, and another candidate is assessed. Over the years, there have been multiple studies looking at ways to improve the trial process even more. This includes studies looking at eyewitness testimony and reliability, similarities between jurors and defendant, and more. For example, research on eyewitness testimony resulted in the creation of a panel of experts called the Technical Working Group for Eyewitness Evidence who published the document "Eyewitness Evidence: A guide for law enforcement" in 1999 to improve police practices and incorporate de-biasing techniques into the criminal justice system (Technical Working Group for Eyewitness Evidence, & United States of America, 1999). Why does all this matter? The purpose of the current study, along with previous research, is to identify potential sources of unfairness that might find their way into the criminal justice system, specifically in court. Collectively, researchers seeking to identify such biases presumably want to make sure

that jury selection and jury instructions given by the judge do not work against the purpose of having a trial -- to have a fair trial.

Past research has shown that jurors tend to be more lenient in sentencing and award amounts when they share similarities with the defendant (Kerr et al., 1995). The examined similarities are typically substantial group similarities such as race, ethnicity, religion, or more. This is known as the similarity-leniency hypothesis (Kerr et al., 1995). Because research on the similarity-leniency hypothesis has primarily examined substantial group similarities, in the present research, I decided to take this a step further and examine non-group, more superficial or trivial similarities. A trivial similarity is a small similarity such as sharing a birthdate, having the same type of fingerprint (loop vs. arch) (Burger et al., 2004), or even dressing similarly, that provides little or no relevant information about a person. Such similarities differ from larger group similarities that many people can share and connect with such as race or gender. Burger et al. (2004) provided research on trivial similarities between participants and a confederate in the realm of compliance with a person's request. They showed that participants who shared a birthdate or a first name with the confederate making a request were more likely to comply with the request (Burger et al., 2004). Miller et al. (1998) used the prisoner's dilemma paradigm to analyze potential effects of trivial similarities. They found that participants who believed they shared a birthday with a partner cooperated more across all types of decisions, including situations that were favorable-to-other (i.e., partner receiving more points than participants for cooperation). They also found that participants showed no difference in ratings of perceived similarity between "birthday-mates" and the control. Thus, both studies showed that trivial similarities can influence the participant, even if they do not explicitly perceive a similarity with the participant. This is an important point for the current research because in a court context, if

trivial similarities have an effect without the juror thinking so, it can lead to biasing in decisions and sentencing which is what this study hopes to solve.

Past research regarding alternative types of trivial similarities investigated the impacts of sharing similar names and the effects of this shared similarity. Burger et al. (2004) found in their study that when participants believed they shared a first name with someone requesting a donation, they ended up donating more money than those who did not share a name. Jones et al. (2004) found in marriage records that people are disproportionately likely to marry someone who has a similar name (e.g., Eric and Erica). This shows that similar effects occur with similar names, not only when the same name is identical. The present research will incorporate the idea of using similar names as one of the trivial similarities that the participant will share. Using similar rather than identical names should reduce the chance of the participant with an uncommon name from noticing the similarity in the study.

Another area of interest for this study regards the effects of shared trivial similarities and liking on juror decisions. Past research has shown that attractive defendants are more likely to receive lighter sentences than unattractive defendants for the same crime, presumably because jurors like attractive defendants more than unattractive defendants (e.g., Landy & Aronson, 1969; Nerneth & Sosis, 1973; Smith & Hed, 1979; Stewart, 1985). If people with whom one shares a trivial similarity are seen as more likeable or socially attractive and more likeable defendants tend to get lighter sentences, might defendants with whom a juror shares a trivial similarity also receive a lighter sentence because they are seen as more likeable? This seems plausible. One study examined a situation in which the participant rated liking of another (fake) participant whose identification number was the same as the birthdate of the real participant. They found that participants reported liking their partner more if their ID number was the same

as their birthdate (Jones et al., 2004). This shows that sharing trivial similarities increases liking. In the context of the current study, this increased liking might lead to the juror being more lenient towards the individual they like and with whom they share trivial similarities.

Although there are many examples of trivial similarities increasing liking and cooperation, past research has also found that such similarities can sometimes actually lead to harsher reactions and negative behaviors. Jiang et al. (2010) investigated the effects of trivial similarities in a sales context. That is, when a participant shared a trivial similarity with someone who was exhibiting an undesirable social behavior, such as being rude to a waitress, the participants distanced themselves and felt less connected to the other person. In the Jiang et al. study, the undesirable behavior led to a more negative attitude towards a program that was being sold by the person exhibiting undesirable behavior and to a lower intention to purchase. Jiang et al. (2010) also found that the possibility of an extended relationship with the undesirable person moderated the negative effects of the shared trivial similarity. When the participant expected future interactions with the person displaying undesirable behavior with whom they share a trivial similarity, they distanced themselves, but when they did not expect any further interactions, they did not distance themselves. This research is relevant to the current study because if participants view the defendant as someone who has engaged in undesirable social behavior (i.e., being in the wrong, or at least engaging in some form of anti-social behavior), then the participant might try to distance themselves and not be lenient in sentencing. Because participants do not expect additional interactions with the defendant, however, perhaps no such distancing would occur.

Present Research

The goal of this study is to determine whether trivial similarities, such as sharing a birthdate, between juror and defendant will affect the verdict and whether jurors will correct for this bias when they are made aware of it. With all the past research in mind about the effects of trivial similarities, my main hypothesis about trivial similarities for the present research is that jurors who share trivial similarities with a defendant will report them as less guilty and will make awards (in the civil case) that are more favorable to the similar defendant.

While there is a significant amount of work examining the biasing effects of trivial similarities, less work has focused on trivial similarities in jury decisions, and how individuals might correct for such biases. To address this, I wanted to examine whether, when made aware of the possibility of a bias, participants will correct for this. Bias can be an unduly favorable or unfavorable reaction toward a person, place, thing, or idea. Beyond similarity or liking towards someone in a trial, there are many features of the defendant that can induce jury bias, including bias towards a defendant's race (e.g., Johnson et al., 1995). To try to eliminate this bias, in trials, before the jury is released to deliberation or to discuss the case to reach a verdict, the judge reads instructions to the jury about the law relevant to the case and about their deliberations. These instructions typically include information about the laws relevant to the case, information that is not to be used in the decision, and other various instructions. Past research has shown that these instructions are, in short, not helpful and often fail to reduce bias (e.g., Ellsworth & Mauro, 1998; Lieberman & Sales, 1997). It is unclear whether the problem lies with the specific instructions that are typically used or with people's motivation or ability to engage in corrections aimed at avoiding bias (Wegener et al., 2000).

With the potential implications for jury trials in the background, it might be helpful to consider past research on bias correction outside the courtroom context. For example, McCaslin

et al. (2010) used a paradigm in which they informed participants that they were "overestimators" or "underestimators" and examined how that labeling changed their later estimates in a dot estimation task. They found that when people were asked to correct for their perceived estimation bias, they corrected in the direction opposite of the label they had been given (i.e., down if they were overestimators and up if they were underestimators). Because participants had been randomly assigned to be labeled as over- or underestimators, the corrections actually made them less accurate in their later dot estimates. In the second experiment, they found that participants became less willing to correct for their supposed bias when they believed there were benefits for the self in having that bias (e.g., if over-estimators are better problem solvers, then there was not as much reason in their minds to correct for an over-estimation bias). Within a jury context, if jurors are aware of the possible bias (based on the trivial similarities) they might over- or under-correct when asked to re-report their verdict. Also, in a jury setting, there is a possibility that they might think there are benefits to keeping their inaccuracies (e.g., "doing justice" or "putting away the bad guy"). Jury trials are supposed to be fair trials and finding a way to make sure people correct for biases and make a fair judgment is one of the key questions that was addressed in this study.

Though not exactly the same as similarity-based liking, some bias correction research investigated corrections of the impacts of source likability in persuasion. One study found that when no correction was given, attitudes toward a proposal were more favorable when the message was presented by a likable rather than dislikable source. But when given a correction instruction, attitudes were more favorable when the message was given by a dislikable rather than likable source (Petty et al., 1998). In a jury context, if a juror finds a defendant to be more likable, possibly due to trivial similarities, they might be more favorable towards them (i.e., less

likely to find them guilty or, if guilty, more lenient in terms of punishment). In the present research, this could mean that if trivial similarities make the defendant seem more likable, the juror might award them more money in a civil judgment. Also, if alerted to the possibility of bias, participants might change their answers, and award more money to a less likable person (sharing no similarities with the participant). Specifically, when made aware of the bias and its possible effects on their initial judgments, participants may attempt to correct for their initial bias and award more money to the other person.

Hypothesis

Taking all the information about similarity-leniency hypothesis, trivial similarities, and bias correction into consideration, my overall hypothesis is that trivial similarities between juror and defendant will influence the juror to be more lenient/award more to the similar defendant than non-similar defendant. When asked to correct for possible bias, however, jurors will correct for their similarity-induced biases, resulting in a reduced bias. With bias correction, there is a possibility of overcorrection which would occur when the participant is made aware of the possible biasing effects and asked to correct for them. Participants might award the non-similar defendant more money/be more lenient as a way to correct the previous bias.

Methods

Participants

Participants were 150 undergraduate psychology students ranging from ages 18 to 37 (M = 19.16 SD = 2.22) who completed the study for class credit. There were 63 (42%) male participants and 87 (58%) female participants who participated.

Procedure

This study used a mixed design where each participant completed both initial and corrected ratings in addition to a between-subject similarity factor. To manipulate similarity, participants were randomly selected to see either a similar defendant or a non-similar defendant (control). After their decisions regarding fault, responsibility, and guilt for the defendant and plaintiff, I asked questions about theories of bias to see if they might predict ratings or shifts in ratings.

After informed consent, participants completed background information for a study regarding jury selection. The general background information about the participant was used to create shared trivial similarities through the use of piping the text into the defendant description. The similarity information included birthdate, not including the year, the participant's specific hometown and the closest major city to their hometown, and the participant's initials. Participants were then told that they had been selected for the jury and would receive information about a civil case. Participants received a police-like report of a description of both defendant and plaintiff adapted from Gleason and Harris (1975). The defendant varied in terms of whether they were similar or not to the participant; the plaintiff did not share any similarities or have specific dissimilarities from the participant.

Materials

Participants saw a fake police summary for the background of the defendant that included name, age, birthdate, height, weight, hair color and hometown (Figure A1). This summary was adapted from Gleason and Harris (1975) and can be found in Appendix A. The name, hometown, and birthdate (not including year) autofilled from the participant answers for background information if selected for the similar condition. Then, the participants saw a fake police description of the plaintiff. In this description, they saw "Name: K.P.", "Age 28 years", "Birth

Date: April 26th, 1992", "Height: 6 feet", "Weight: 183 lbs.", "Hair Color: Brown", and "Hometown: Omaha, Nebraska". None of the information included in the description of the plaintiff changed across participants, only the defendant information changed across participants within the similar condition.

Participants then received a brief case summary of a minor fender bender that was also adapted from Gleason and Harris (1975). For the full summary, see Figure A2 in Appendix A. The case described when the accident occurred, where it occurred, and a detailed description of the scene. It included statements from both the defendant and plaintiff about the accident, a statement about police estimation of damage, and an eyewitness report. The case was somewhat ambiguous regarding who was at fault to enable potential similarity-based biases to influence their uncorrected judgments. After the descriptions, there was a reading check to make sure that participants paid attention to the information. These were multiple choice questions asking where the plaintiff's hometown was and when the accident occurred.

Finally, the participants were given instructions based on the information given. They were asked to report guilt, responsibility, and fault of both the plaintiff and defendant using a 5-point scale ranging from "Not <guilty/at fault/responsible at all>" to "Very <guilty/much at fault/responsible>". Then, they were told that both the plaintiff and defendant are willing to settle, and the participant was to divide $10,000 between them for damages. This measure was set up using a toggle bar to distribute money to both plaintiff and defendant where the total had to add up to $10,000. After the initial decisions, regardless of condition, the participants were told that while they were completing the questions, a possible bias was noticed between them and the defendant and the researchers would like for them to re-do their answers taking into account that potential bias. Participants were then asked to re-report their answers of guilt,

responsibility, and fault, and asked again to re-distribute money to both plaintiff and defendant using the same measures as they used at time 1. Participants were then thanked for their participation and debriefed about the purpose of the study.

Results

Manipulation Check

A manipulation check was done to investigate whether the manipulation of similarity condition affected liking of the defendant. Results indicated that the manipulation of similarity condition did significantly affect how much the defendant was liked $t(148) = 2.426$, $p = 0.016$. The similar defendant was liked more ($M = 3.087$, $SD = 0.487$) than the non-similar defendant ($M = 2.880$, $SD = 0.555$).

Guilt

To analyze the impact of Similarity condition (similar vs. non-similar) on guilt prior to correction instructions, I conducted an independent samples t-test on ratings of defendant guilt. The effect of Similarity on uncorrected guilt was not significant, $t(148) = 0.404$, $p = 0.687$. Even if similarity did not influence uncorrected defendant guilt, correction instructions could prompt shifts in judgments. To examine this possibility, I conducted a mixed-design general linear model analysis that included the Similarity condition, the Time at which ratings were made (pre- vs post-correction) and the interaction of those two factors. Results showed no significant main effect of defendant Similarity, $F(1, 148) = 0.495$, $p = 0.483$. There was also no significant main effect for the Time at which guilt ratings were taken (pre- vs post-correction-instruction), $F(1, 148) = .053$, $p = 0.818$. Finally, the 2-way interaction was also not significant ($F[1,148]= 0.214$, $p=0.645$). Please see Table B1 in Appendix B for means and standard deviations of guilt ratings

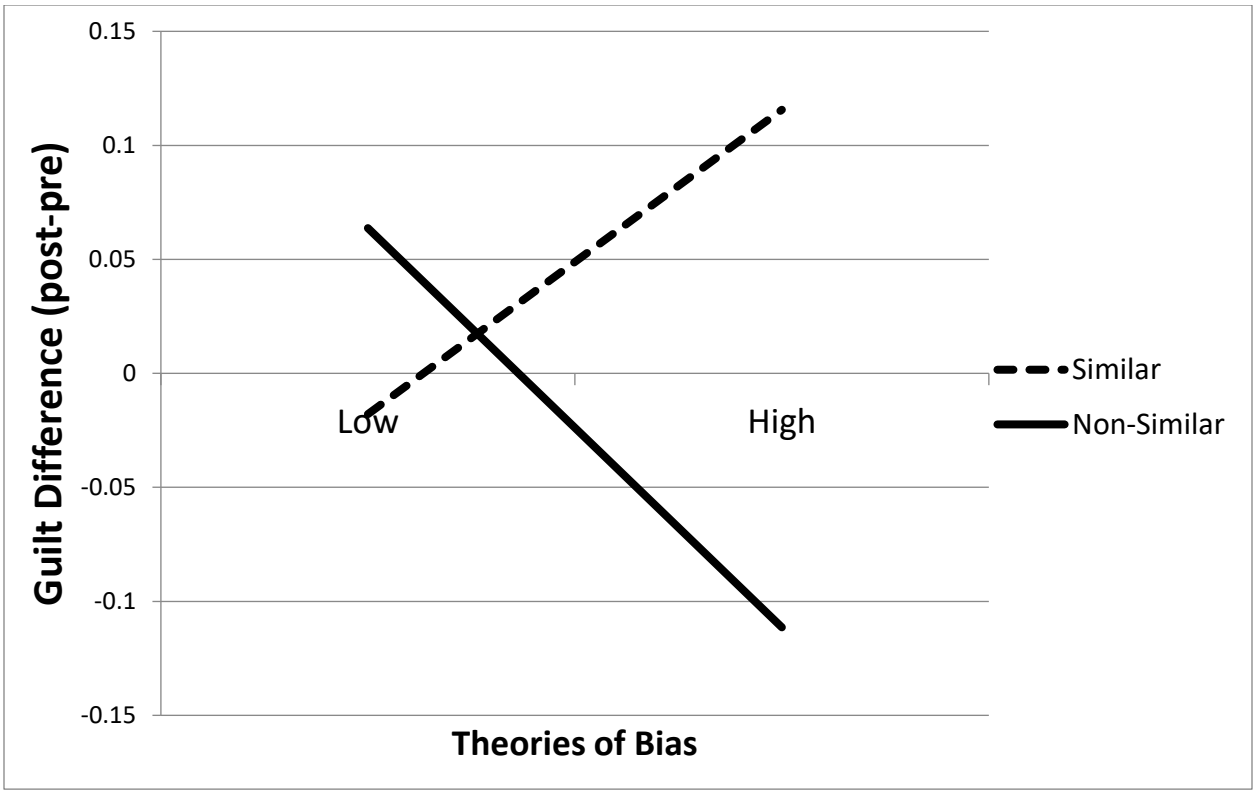
as a function of Time and Similarity and Table B2 for means and standard deviations of the difference in guilt ratings as a function of Similarity.

Perhaps there were not overall shifts in ratings from pre- to post-correction-instructions because participants simply did not view similarity as likely to affect initial perceptions. To examine this possibility, a single-sample t-test was used to analyze participant's perceptions of being affected by similarity where answers were compared against the midpoint value which was a neutral response. Results were significant $t(149) = 15.354$, $p < 0.0001$ and show that participants did believe that similarity with a defendant would be associated with leniency (M = 6.674, SD = 1.335). Considering that results indicate participants believed similarity would influence their opinion, the lack of significant results might be due to the manipulations being too subtle or too noticeable, which is discussed further later. Although participants believe that their perceptions would be affected by similarity, I investigated whether theories of bias might predict ratings. I examined the overall reported theories of bias and used them to predict shifts from pre- to post-correction-instruction ratings (post-correction ratings minus pre-correction ratings). The theories of bias measures were used to address how participants thought similarity might influence guilt, fault, and responsibility ratings, regardless of the condition they were in. For example, one theories of bias question asked, "Do you think that sharing a name or birthdate with someone would affect your opinions of that individual?" where they rated their answer on a 9-point scale ranging from "Opinions would be less favorable towards the individual I like" to "Opinions would be more favorable towards the individual I like". A regression analysis was then completed to investigate how theories of bias might moderate the impact of condition on difference in guilt perceptions of the defendant. The key interaction between theories of bias and similarity condition was not significant, $F(1, 149) = 1.14$, $p = 0.336$, and can be seen in Figure 1

below. In follow-up analyses to investigate the simple effects of theories of bias within each similarity condition, theories of bias did not significantly predict guilt difference for those in the similar condition, B=0.067, $t(74) = 1.094$, $p = 0.278$, or for those in the non-similar condition, B=-0.088, $t(74) = -1.432$, $p = 0.156$. Though none of these results were significant, when theories of bias were high, those in the similar condition reported directionally higher guilt ratings than those in the non-similar condition, but when theories of bias were low, those in the similar condition reported lower guilt ratings. Put differently, for those in the similar condition, post-correction guilt ratings were higher than pre-correction ratings when theories of bias were high versus when theories of bias were low. For those in the non-similar condition, the opposite occurred; post-correction guilt ratings were lower than pre-correction ratings when theories of bias were high versus when theories of bias were low. This pattern was directionally consistent with my hypothesis, albeit not significantly so.

Figure 1

Condition by Theories of Bias on Guilt



Responsibility

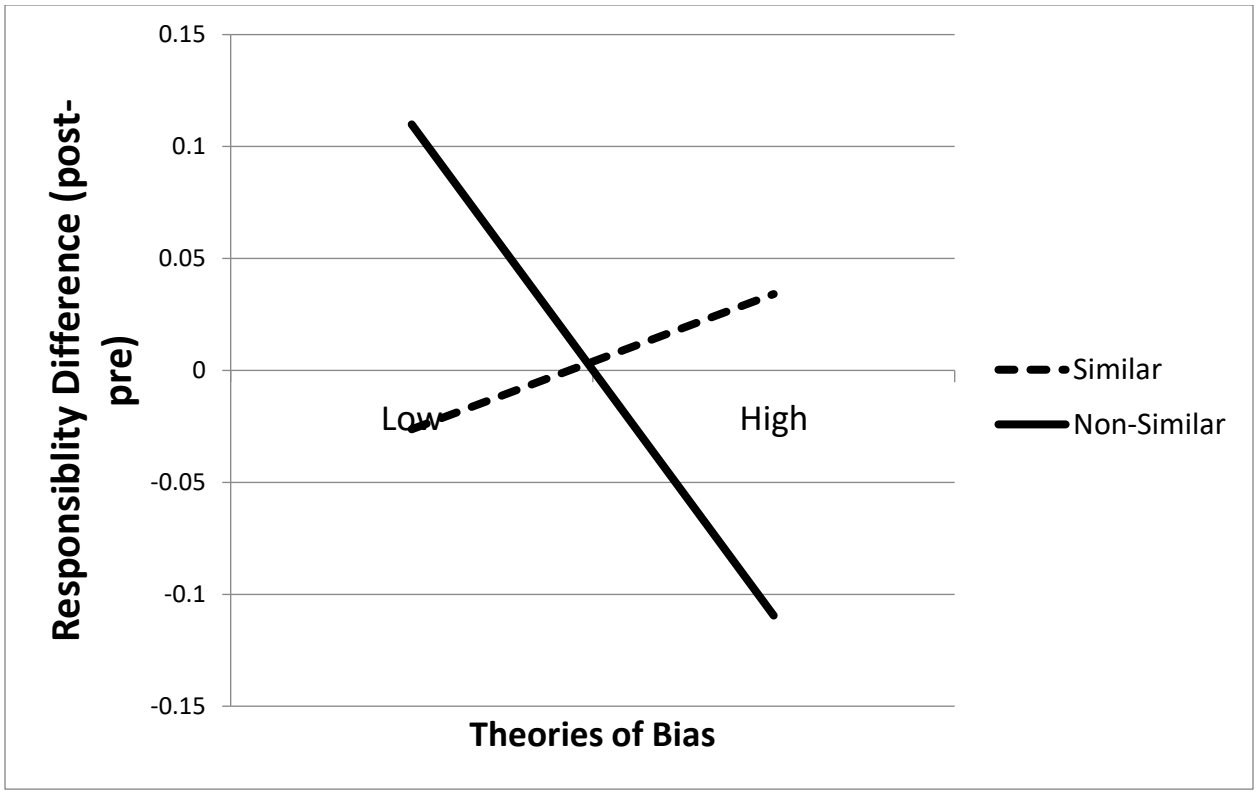
To analyze the impact of Similarity condition (similar vs. non-similar) on responsibility prior to correction instructions, I conducted an independent samples t-test on ratings of defendant responsibility. The effect of Similarity on uncorrected responsibility was not significant, $t(147) = 1.018$, $p = 0.310$. Even if similarity did not influence uncorrected defendant responsibility, correction instructions could prompt shifts in judgments. To examine this possibility, I conducted a mixed-design general linear model analysis that included the Similarity condition, the Time at which ratings were made (pre- vs post-correction) and the interaction of those two factors. Results showed no significant main effect of defendant Similarity, $F(1,147) = 1.388$, $p = 0.241$. There was also no significant main effect for the Time at which responsibility ratings were taken (pre- vs post-correction-instruction), $F(1,147) = 0.007$, $p = 0.933$. Finally, the 2-way interaction was also not significant, $F(1,147) = 0.007$, $p = 0.933$. Please see Table B1 in Appendix B for

means and standard deviations of responsibility ratings as a function of Time and Similarity and Table B2 for means and standard deviations of the difference in responsibility ratings as a function of Similarity.

Again, I examined the overall reported theories of bias and used them to predict shifts from pre- to post-correction-instruction ratings to possibly explain why there were no overall shifts in ratings from pre- to post-correction instructions. A regression analysis was then conducted to investigate how theories of bias might moderate the impact of similarity condition on the difference in responsibility perceptions of the defendant. Results for the theories of bias by similarity interaction were not significant, $F(1,148) = 0.576$, $p = 0.632$, and can be seen below in Figure 2. In follow-up analyses to investigate the simple effects of theories of bias within each similarity condition, theories of bias did not significantly predict responsibility shifts for those in the similar condition, $B = 0.030$, $t(73) = 0.400$, $p = 0.691$, or for those in the non-similar condition, $B = -0.110$, $t(74) = -1.174$, $p = 0.244$. For those in the similar condition, post-correction responsibility ratings are higher than pre-correction ratings when theories of bias are high versus when theories of bias are low. For those in the non-similar condition, post-correction responsibility ratings are lower than pre-correction ratings when theories of bias are high versus when theories of bias are low. This is similar to what was seen in Figure 1 for guilt ratings but was again quite weak.

Figure 2

Condition by Theories of Bias on Responsibility



Fault

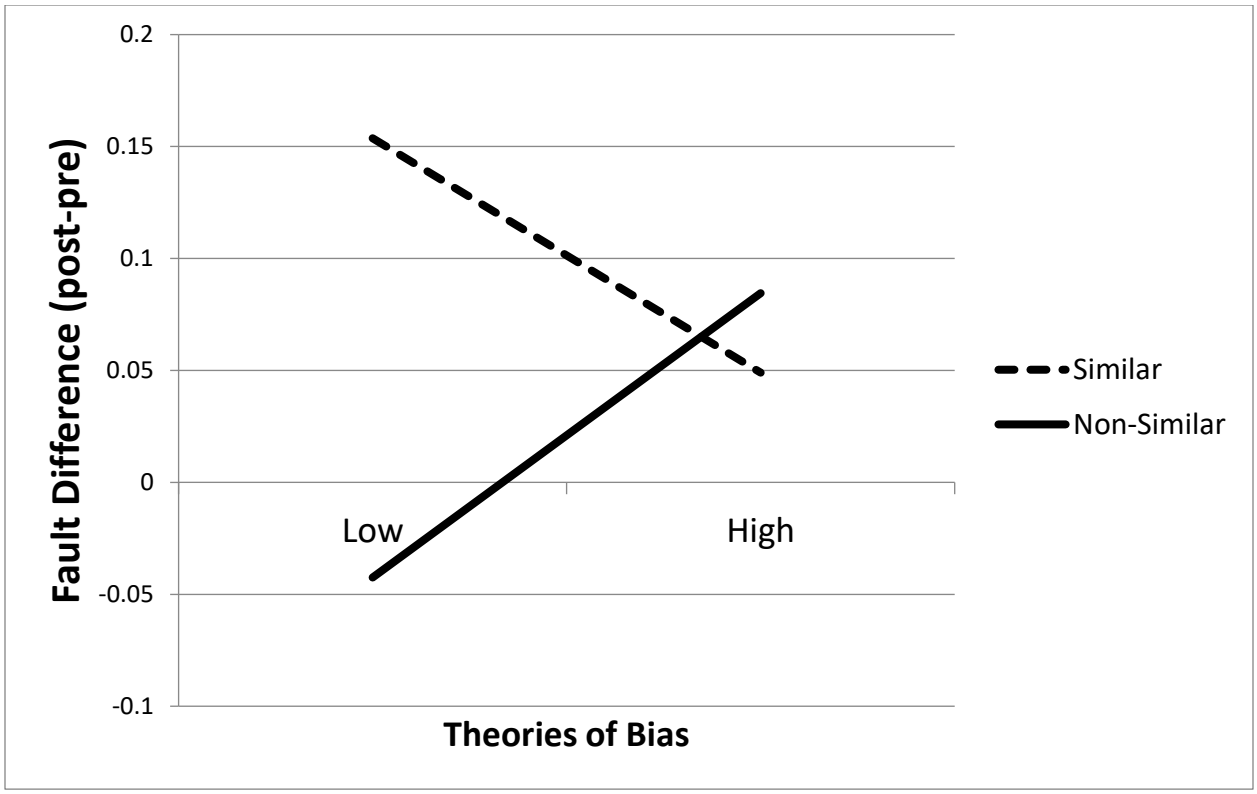
To analyze the impact of Similarity condition (similar vs. non-similar) on fault prior to correction instructions, I conducted an independent samples t-test on ratings of defendant fault. The effect of Similarity on uncorrected fault was not significant. $t(147) = 0.331$, $p = 0.741$. Even if similarity did not influence uncorrected defendant fault, correction instructions could prompt shifts in judgments. To examine this possibility, I conducted a mixed-design general linear model analysis that included the Similarity condition, the Time at which ratings were made (pre- vs. post-correction) and the interaction of those two factors. Results showed no significant main effect of defendant Similarity, $F(1,147) = 0.621$, $p = 0.432$. There was also no significant main effect for the Time at which fault ratings were taken (pre- vs post-correction-instruction), $F(1,147) = 0.896$, $p = 0.345$., Finally, the 2-way interaction was also not significant, $F(1,147) = 0.546$, $p = 0.461$. Please see Table B1 in Appendix B for means and standard deviations of fault

ratings as a function of Time and Similarity and Table B2 for means and standard deviations of the difference in fault ratings as a function of Similarity.

Again, I examined the overall reported theories of bias and used them to predict shifts from pre- to post-correction-instruction ratings to possibly explain why there were no overall shifts in ratings from pre- to post-correction instructions. A regression analysis was then conducted to investigate how theories of bias might moderate the impact of similarity condition on fault perceptions of the defendant. Results of the interaction of theory and similarity were not significant, $F(1,148) = 0.663$, $p = 0.576$ and can be seen below in Figure 3. In follow-up analyses to investigate the simple effects of theories of bias within each similarity condition, theories of bias did not significantly predict fault difference for those in the similar condition, $B = -0.052$, $t(73) = -0.806$, $p = 0.423$, or for those in the non-similar condition, $B = 0.064$, $t(74) = 0.889$, $p = 0.377$. The fault ratings showed the opposite pattern of what was seen for the guilt ratings. For those in the similar condition, post-correction fault ratings were higher than pre-correction ratings when theories of bias were low versus when theories of bias were high. For those in the non-similar condition, post-correction fault ratings were lower than pre-correction ratings when theories of bias were low versus when theories of bias were high.

Figure 3

Condition by Theories of Bias on Fault



Money

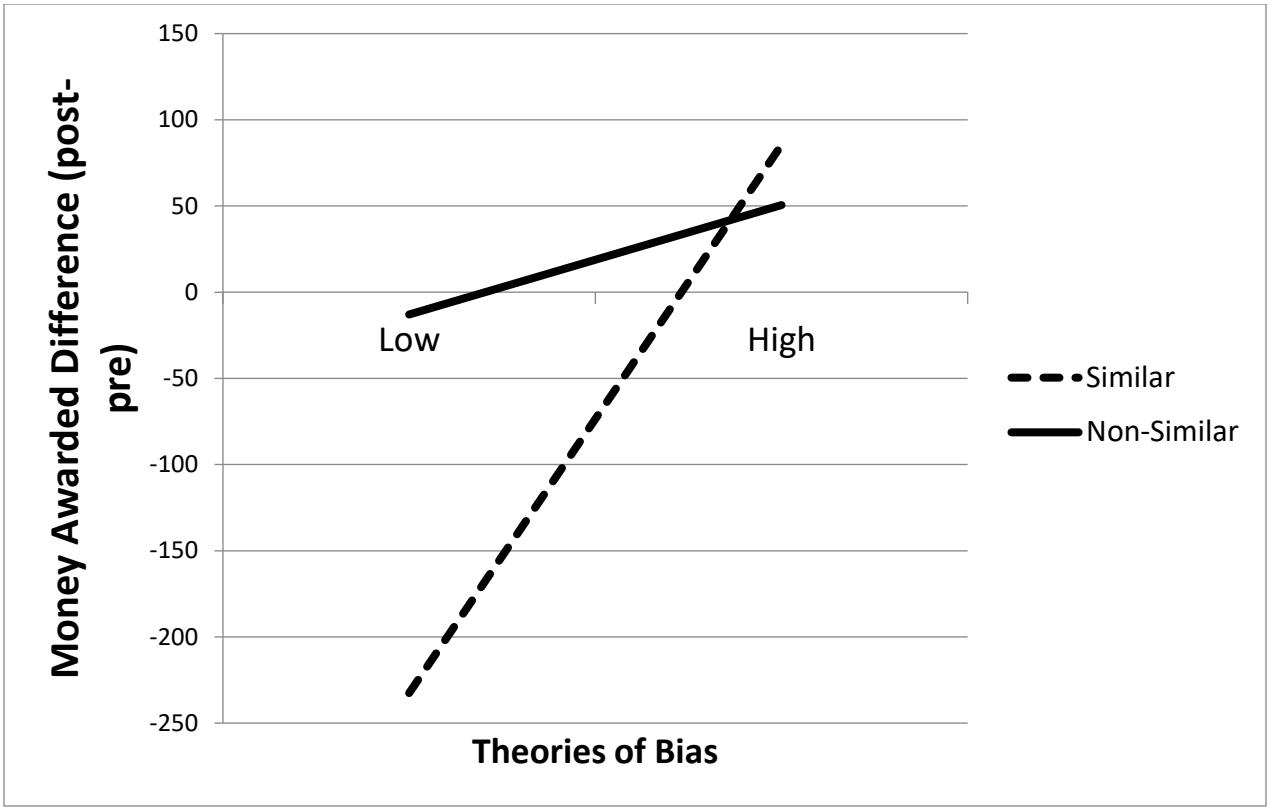
To analyze the impact of Similarity condition (similar vs. non-similar) on money awarded prior to correction instructions, I conducted an independent samples t-test on the amount of money awarded to the defendant. The effect of Similarity on uncorrected money awarded was not significant. $t(149) = -0.293$, $p = 0.699$. Even if similarity did not influence uncorrected defendant money awarded, correction instructions could prompt shifts in judgments. To examine this possibility, I conducted a mixed-design general linear model analysis that included the Similarity condition, the Time at which ratings were made (pre- vs. post-correction) and the interaction of those two factors. Results showed no significant main effect of defendant Similarity, $F(1,149) = 0.273$, $p = 0.602$. There was also no significant main effect for the Time at which amount of money awarded ratings were taken (pre- vs post-correction-instruction), $F(1,149) = 0.135$, $p = 0.714$. Finally, the 2-way interaction was also not significant, $F(1,149) =$

0.086, $p = 0.770$. Please see Table B1 in Appendix B for means and standard deviations of money awarded as a function of Time and Similarity and Table B2 for means and standard deviations of the difference in money awarded as a function of Similarity.

Again, I examined the overall reported theories of bias and used them to predict shifts from pre- to post-correction-instruction ratings to possibly explain why there were no overall shifts in ratings from pre- to post-correction instructions. A regression analysis was then conducted to investigate how theories of bias might moderate the impact of similarity condition on money awarded to the defendant. Results of the interaction of theory and similarity were not significant, $F(1,149) = 0.442$, $p = 0.724$ and can be seen below in Figure 4. In follow-up analyses to investigate the simple effects of theories of bias within each similarity condition, theories of bias did not significantly predict the difference in money awarded for those in the similar condition, $B = 159.013$, $t(74) = 1.077$, $p = 0.285$, or for those in the non-similar condition, $B = 31.741$, $t(74) = 0.199$, $p = 0.843$. For those in the similar condition, post-correction money awarded ratings were lower than pre-correction ratings when theories of bias were low versus when theories of bias were high. For those in the non-similar condition, post-correction money awarded ratings were also lower than pre-correction ratings when theories of bias were low versus when theories of bias were high.

Figure 4

Condition by Theories of Bias on Money Awarded



General Discussion

Although this study produced no overall significant effects, results were trending in the direction that was predicted for two of the three dependent variables. My hypothesis was that trivial similarities between juror and defendant matter and will influence the juror to be more lenient/award more to the similar defendant than non-similar plaintiff and when asked to correct for possible bias, jurors should correct, resulting in a reduced bias. Results indicated that those who were in the similar condition tended only slightly to report higher guilt and responsibility ratings post-correction. That direction was consistent with my hypothesis that when given correction instructions, participants will correct more when they share trivial similarities with the defendant. These results were not replicated for fault ratings, but I will discuss possible reasons why that occurred later in this section. Results also showed that the manipulations were

significant in affecting a change in liking of the defendant, but not in ratings of guilt, fault, responsibility, and money awarded. This could mean that maybe participants noticed enough of the similarity manipulation to like the similar defendant more, but it was not enough to affect their ratings of the defendant possibly because the manipulation was not effective enough. Other possibilities of why this occurred will be discussed later in this section.

Beyond the main effects of similarity for guilt and responsibility, I also found that the higher the theories of bias, the larger the difference between guilt ratings based on condition. Specifically, there was a larger difference between guilt ratings when theories of bias were high compared to when theories of bias were low. This was also seen for responsibility ratings as well, but the difference between responsibility ratings when theories of bias were high is not as large as the difference in guilt ratings when theories of bias were high. For fault ratings, the opposite pattern was seen. When theories of bias were low, the difference in fault ratings was much larger than when theories of bias were high. It is unknown why this occurred only for fault ratings but it is an interesting aspect of the study results. This could possibly be due to the fact that, because the case was ambiguous, participants might not have felt comfortable assigning fault, or they may be set on who was a fault and not want to change their answer.

If significant results had been found, this would have had implications for courts in both the judge's instruction and the jury selection process. The instructions that judges give to jurors before they go into deliberation are in need of improvement overall, but if results were significant, issues regarding the effects of similarities with plaintiffs or witnesses for both the prosecution and defense would need to be examined and addressed to take steps toward improvement suggestions to judge instructions. Also, if significant results were found, there would be implications for jury selection and voir dire (process of selecting jurors for a trial). If

significant, it could mean that lawyers might be able to select or exclude people for a jury based on trivial similarities they may share with either the defendant or plaintiff, and these similarities could be a factor in the overall voir dire process.

Limitations and Future Directions

Noticeable Similarity

In this study, my goal was to make the shared similarities subtle enough to possibly create a bias between the participant and defendant. In doing this, I may have pushed the boundary a bit too far to be subtle, which may have led participants to notice the similarities and adapting their ratings to this information. In the study, to introduce the subtle similarities in the police description of the defendant, the information that each participant put in for their background questions was auto-filled into the initials, hometown, and birthdate for the defendant. If the participant misspelled anything or formatted their answer oddly, it would autofill verbatim, which might have led to them noticing the similarities and reacting differently. For example, if they mis-typed or had odd formatting, they might have been more likely to notice the manipulation and change their answers to what they think the research might have expected (or tried not to be affected by the similarity because it seemed forced on them). Since participants did believe that similarity would have an impact on their ratings, by noticing the results they might have perceived the defendant as not truly being similar to them, but as being "faked" because of the piped information and odd formatting. Future studies could adapt to this limitation by possibly giving a pre-test at an earlier date to collect background information, and then create unique studies for each participant that are given at a later date, using the information they provided previously with properly formatted information for the defendant to eliminate formatting errors that are easily noticed.

Subtle Similarity

Another significant limitation is that due to the COVID-19 pandemic, participants were not being brought into the lab and were taking the study elsewhere (i.e., at home) which didn't allow us as researchers to control the environment where the participant would take the study. This may have affected the results that were obtained because the environment they took the study in could not be controlled and could lead to distractions that kept them from paying enough attention to the study and therefore might have missed similarity manipulations. Bringing participants into the lab allows for the testing environment to be controlled and distraction free, whereas taking the study at home, there are multiple distraction possibilities that could lead the participant to not pay attention to the study details and miss important information in the study. Being able to control the testing environment eliminates these distractions and possible future replications post-pandemic could examine whether this was a significant determinant of the results obtained.

Another reason why this study may have lacked significant results was that the manipulations might have been too weak given the available sample size. Although I was trying to make the similarities subtle, there is a possibility that they were too subtle, and participants did not notice enough relevant information to create bias. Since results indicated that participants viewed similarity as having an impact on their perceptions, having under-powered manipulations might not have created enough similarity to have a significant biasing effect. Future studies could examine the possibility of adding more trivial similarities to the study to strengthen the manipulation or find other methods to increase power to investigate whether a higher-powered version of the study might yield significant results.

Negative Similarity Effects

If my measures were too subtle or too obvious, future studies could investigate whether there are conditions under which trivial similarities no longer have an effect on juror opinions or if they even might sometimes have a negative effect. In Jiang et al. (2010), they found that when participants share trivial similarities with someone displaying negative characteristics, they try to distance themselves and see the person as less likeable. It would be interesting to see if this could happen in a courtroom context and if the similarities can be subtle enough to create bias, but not too obvious to lead to unintended effects. For example, if jurors share trivial similarities with a defendant who either is displaying negative characteristics, either in the courtroom during the trial (e.g., screaming at judge, swearing at lawyer/judge), or in the crime they committed (e.g., violent crime instead of civil crime) would this lead the juror to try to distance themselves, and at what point would that happen?

Future Study Changes

Because there were no significant results in the current study, there are quite a few next steps that could be taken to revise the study to possibly find more informative results. Due to time constraints after the IRB protocol was approved, I was unable to run pre-tests for the measures and was unable to edit the study and test participants again. One of the first things I would edit would be the measures for guilt, fault, and responsibility. All of these measures address whether the participant is more lenient towards the defendant, but I think that in the future, an expanded set of guilt, fault, and responsibility measures should be included so that a more detailed analysis can be made to see if a possible bias occurred. Another item I would change is the way that the study was distributed. Initially, I attempted to run a 2-part study where initial judgments were followed up at a later date from the original judgments, but retention was a major limitation that led to the use of a full single-session study instead. Using this two-part

format for the study would better simulate a typical court proceeding where the correction instructions come at a later time than the initial information.

Conclusion

This study investigated the potential effects of trivial similarities shared between juror and defendant and whether these biases could be corrected after encountering bias correction instructions. Overall, results were not significant but patterns on two of the outcome measures were trending in the direction of my hypothesis. Although this study did not find significant results, the results that were found pave the path for future steps and studies that investigate the impact of trivial similarities in a courtroom context and have implications for judge instructions and jury selection in trials.

References

Burger, J. M., Messian, N., Patel, S., Del Prado, A., & Anderson, C. (2004). What a coincidence! The effects of incidental similarity on compliance. *Personality and Social Psychology Bulletin*, *30*(1), 35-43.

Cong. Rsch. Serv. (n.d.). Right to an Impartial Jury: Current Doctrine. In *Constitution of the United States: Analysis and Interpretation*. Retrieved March 22, 2021, from https://constitution.congress.gov/browse/essay/amdt6-3-2-1-1/ALDE_00000940/

Cong. Rsch. Serv. (n.d.). Right to Trial by Jury: Scope of the Right. In *Constitution of the United States: Analysis and Interpretation*. Retrieved March 22, 2021, from https://constitution.congress.gov/browse/essay/amdt6_3_1_2_2/

Ellsworth, P. C., & Mauro, R. (1998). Psychology and law. *The handbook of social psychology*, *2*, 684-732

Gleason, J. M., & Harris, V. A. (1975). Race, socio-economic status, and perceived similarity as determinants of judgements by simulated jurors. *Social Behavior and Personality: an international journal*, *3*(2), 175-180.

Jiang, L., Hoegg, J., Dahl, D. W., & Chattopadhyay, A. (2010). The persuasive role of incidental similarity on attitudes and purchase intentions in a sales context. *Journal of Consumer Research*, *36*(5), 778-791.

Johnson, J. D., Whitestone, E., Jackson, L. A., & Gatto, L. (1995). Justice is still not colorblind: Differential racial effects of exposure to inadmissible evidence. *Personality and Social Psychology Bulletin*, *21*(9), 893-898.

Jones, J. T., Pelham, B. W., Carvallo, M., & Mirenberg, M. C. (2004). How do I love thee? Let me count the Js: implicit egotism and interpersonal attraction. *Journal of personality and social psychology*, *87*(5), 665.

Kerr, N. L., Hymes, R. W., Anderson, A. B., & Weathers, J. E. (1995). Defendant-juror similarity and mock joror judgments. *Law and Human Behavior*, *19*(6), 545-567.

Landy, D., & Aronson, E. (1969). The influence of the character of the criminal and his victim on the decisions of simulated jurors. *Journal of Experimental Social Psychology*, *5*(2), 141-152.

Lieberman, J. D., & Sales, B. D. (1997). What social science teaches us about the jury instruction process. *Psychology, Public Policy, and Law*, *3*(4), 589.

McCaslin, M. J., Petty, R. E., & Wegener, D. T. (2010). Self-enhancement and theory-based correction processes. *Journal of Experimental Social Psychology*, *46*(5), 830-835.

Miller, D. T., Downs, J. S., & Prentice, D. A. (1998). Minimal conditions for the creation of a unit relationship: The social bond between birthdaymates. *European Journal of Social Psychology*, *28*(3), 475-481.

Nemeth, C., & Sosis, R. H. (1973). A simulated jury study: Characteristics of the defendant and the jurors. *The Journal of Social Psychology*, *90*(2), 221-229.

Petty, R. E., & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology*, *29*(2), 137-165.

Petty, R. E., Wegener, D. T., & White, P. H. (1998). Flexible correction processes in social judgment: Implications for persuasion. *Social cognition*, *16*(1), 93-113.

Smith, E. D., & Hed, A. (1979). Effects of offenders' age and attractiveness on sentencing by mock juries. *Psychological Reports*, *44*(3), 691-694.

Stewart, J. E. (1985). Appearance and punishment: The attraction-leniency effect in the courtroom. *The Journal of social psychology*, *125*(3), 373-378.

Technical Working Group for Eyewitness Evidence, & United States of America. (1999). Eyewitness Evidence: A Guide for Law Enforcement.

Wegener, D. T., Kerr, N. L., Fleming, M. A., & Petty, R. E. (2000). Flexible corrections of juror judgments: Implications for jury instructions. *Psychology, Public Policy, and Law*, *6*(3), 629.

Appendix A

Research Materials

Figure A1

Police Summary of Defendant

SUMMARY OF POLICE FILES ON BACKGROUND OF THE DEFENDANT

NAME: T.T. – will change to same initials as the participant which was listed in background questionnaire.

AGE: 26 years

BIRTH DATE: March 3, 1994 – will change to same birthdate participant listed in background questionnaire, but not birth year.

HEIGHT: 6 feet, 2 inches

WEIGHT: 170 lb.

HAIR COLOR: Brown

HOMETOWN: Cleveland, Ohio – will change to major city participant has listed in background questionnaire.

Figure A2

Police Summary of Plaintiff

SUMMARY OF POLICE FILES ON BACKGROUND OF THE PLAINTIFF

NAME: K.P.

AGE: 28 years

BIRTH DATE: April 26, 1992

HEIGHT: 6 feet

WEIGHT: 183 lb.

HAIR COLOR: Brown

HOMETOWN: Omaha, Nebraska

Figure A3

Case Summary

SUMMARY OF CASE:

The defendant and plaintiff were involved in a motor vehicle accident on Tuesday, April 18th, around 4:45pm. The accident occurred on the corner of Broad St. and Elm St. at an intersection. The defendant was driving northbound on Elm while the plaintiff was turning right from Broad to Elm, when the two cars made contact. The plaintiff states s/he looked both ways multiple times before making the turn. The defendant states that s/he did not see a car at the intersection and proceeded through. Police estimate that the damage to the defendant/plaintiff's car was significantly worse than the damage to the other. An eyewitness reported that they "…didn't see the whole thing, but it looks like the (defendant/plaintiff) was in the wrong".

Appendix B

Tables of Means and Standard Deviations

Table B1

Means and Standard Deviations for Ratings Towards Defendant as a Function of Both Time and Condition

	Pre-correction		Post-correction	
Condition	*M*	*SD*	*M*	*SD*
		Guilt		
Similar	3.6	0.788	3.64	0.69
Non-similar	3.55	0.827	3.53	0.811
		Fault		
Similar	3.54	0.863	3.65	0.814
Non-similar	3.49	0.876	3.51	0.76
		Responsibility		
Similar	3.54	0.814	3.53	0.759
Non-similar	3.4	0.87	3.41	0.931
		Money		
Similar	4819.59	2026.96	4727.91	2153.25
Non-Similar	4936.15	2116.2	4925.51	1946.09

Table B2

Means and Standard Deviations for Difference in Ratings of Defendant as a Function of Condition

	Shift/Difference (post - pre)	
Condition	*M*	*SD*
	Guilt	
Similar	0.04	0.725

Non-Similar	-0.013	0.688
	Fault	
Similar	0.108	0.769
Non-Similar	0.013	0.797
	Responsibility	
Similar	0	0.891
Non-Similar	0.013	1.046
	Money	
Similar	-94.68	1753.56
Non-Similar	-10.63	1772.89